Skill of the MJO and Northern Hemisphere Blocking in GEFS Medium-Range Reforecasts

Thomas M. Hamill and George N. Kiladis

NOAA Earth System Research Lab
Physical Sciences Division
Boulder, Colorado, USA

Submitted to *Monthly Weather Review*

14 June 2013

Corresponding author address:

Dr. Thomas M. Hamill
NOAA Earth System Research Lab, Physical Sciences Division
R/PSD1, 325 Broadway
Boulder, Colorado USA 80305-3328
e-mail: tom.hamill@noaa.gov
phone: +1 (303) 497-3060 telefax +1 (303) 497-6449

ABSTRACT

Forecast characteristics of Northern Hemisphere atmospheric blocking and the MJO were diagnosed using an extensive time series (Dec-Jan-Feb 1985-2012) of daily medium-range ensemble reforecasts based on an approximately stable version of the NCEP Global Ensemble Forecast System (GEFS).

For blocking : (a) the GEFS slightly under-forecasted blocking frequency at longer leads in the Euro-Atlantic sector, and inter-annual variability of blocking frequency was quite large; (b) predictive skill of actual blocking was substantially smaller than its perfect-model skill; (c) block onset and cessation were forecast less well than overall blocking; (d) there was substantial variability of blocking skill between half-decadal periods; and (e) the reliability of probabilistic blocking forecasts degraded with increasing lead time.

For the MJO: (a) forecasts of strong MJOs propagated too slowly, especially the component associated with outgoing longwave radiation (OLR), i.e., convection; (b) tropical precipitation was greatly over-forecast at early lead times; (c) the ensemble predictions were biased and/or under-dispersive, manifested in U-shaped rank histograms of MJO indices. Magnitude forecasts were especially U-shaped. (d) bi-variate MJO correlation skill was larger for its wind than for its OLR component, and was larger for the higher-amplitude MJO events; (e) there was some half-decadal variability in skill; (f) probabilistic skill of the MJO forecast was modest, and skill was larger when measured relative to climatology than when measured relative to a lagged persistence forecast.

63 For longer-lead forecasts, the GEFS demonstrated little ability to replicate the
64 changes in blocking frequency due to a strong MJO that were noted in analyzed data.
65

1. **Introduction**.

Medium-range predictability and forecastability[1] of low-frequency modes of atmospheric variability can be evaluated more readily with a long time series of forecasts. In this manuscript we evaluate two such modes of variability, Northern Hemispheric atmospheric blocking and the Madden-Julian Oscillation (MJO; Madden and Julian 1971) and their interactions using an extensive set of global medium-range ensemble reforecasts.

Both the MJO and blocking occur somewhat infrequently at a particular location. Using the blocking definition of Tibaldi and Molteni (1990), Northern Hemisphere wintertime blocking occurs approximately 2% to 22% of the time, averaged over several decades. Blocks are most common over the eastern Atlantic Ocean and western Europe, with a secondary frequency maximum in the central Pacific Ocean. Blocks can be persistent, leading to both long periods of rather similar, and often high-impact weather. For the slowly moving, large-scale, equatorially trapped tropical convective clusters and associated wind perturbations known as the MJO, a given year may produce only a couple coherent, high-amplitude MJO events, although this number will vary depending on the definition of the MJO (Straub 2013).

[1] *Predictability* here refers to an intrinsic property of the atmospheric process and measures the time scale at which two initially similar but not identical perturbed initial conditions will become as different as random draws of atmospheric states. While numerical models are often used to estimate the (unknown) predictability, the predictability is not a forecast property. *Forecastability*, in contrast, indicates the ability of the forecast model to provide guidance that the user will judge to have some value. Forecastability may be evaluated with many different metrics.

Evaluating the predictability and forecastability of blocking and the MJO are challenging given the infrequency and the temporal continuity of these events, the latter of which reduces the effective sample size (Wilks 2006, eq. 5.12). If one is interested, for example, in the ability of the model to forecast block onset for a particular longitude band, a three-month period may have, say, 15 days with blocked conditions, but most likely those 15 days occurred in 1-3 persistent blocking events. Similar issues occur with the MJO. The interaction of two such events is even more difficult to evaluate with limited samples, e.g., evaluating the change in blocking frequency related to a large forecast MJO event in the Indian Ocean. A large sample size could be provided from a reforecast, i.e., a multi-year or preferably a multi-decadal sample of forecasts from a fixed forecast model and assimilation system.

Extensive reforecasts (hindcasts) often haven't been available to facilitate such studies. For blocking, the most comprehensive recent study was by Jung et al. (2012), which used hindcasts to document blocking frequency in extended-range simulations from the European Centre for Medium-Range Weather Forecasts (ECMWF) model at various resolutions. The authors found that Euro-Atlantic sector blocking frequency was generally more under-forecast with lower-resolution models. Older studies included the Watson and Colucci (2002) study of Northern Hemispheric wintertime blocking using data from the operational NCEP global spectral model from 1995-1998, and Mauritsen and Källén (2004), who studied blocking in the ECMWF system during the Northern Hemispheric 2000-2001 winter. Both studies found too few blocks in the forecast. Pelly and Hoskins (2003a), using

107 a potential vorticity-based method of defining blocks, evaluated the ECMWF output
108 for a year of data beginning on 1 August 2001, a period spanning several model
109 changes. Though they also found blocking frequency was under-forecast, there was
110 positive skill in the probabilistic forecasts of blocks out to 10 days, and they found
111 block onset was better forecast than block cessation.

112 There have been more recent studies of the MJO than for blocking, for
113 seasonal simulations at least. The MJO is being actively studied in part because it
114 affects monsoon (Yasunari 1979) and tropical cyclone variability (Maloney and
115 Hartmann 2000ab). The MJO also can excite extra-tropical Rossby wave trains
116 (Knutson and Weickmann 1987, Jones et al. 2004, Weickmann and Berry 2009) and
117 can interact with mid-latitude, low-frequency modes of variability such as the North
118 Atlantic Oscillation (Lin et al. 2009). The MJO is often poorly forecast, and there is
119 some evidence that an improved MJO forecast may result in improved mid-latitude
120 forecasts (Ferranti et al. 1990, Vitart and Molteni 2010).

121 The interaction of the MJO with the mid-latitude flow and its forecastability
122 has been an active area of investigation (Leibmann and Hartmann 1984, Kiladis and
123 Weickmann 1992, Hendon et al. 2000; Riddle et al. 2012). Since the MJO represents
124 the variability at time scales of 30-70 d^{-1}, (e.g., Waliser et al. 2009), it has been
125 common to examine numerical simulations to lead times of several months and to
126 leverage hindcasts to provide large-enough samples. Examples of studies with
127 seasonal forecast models and hindcasts include Hendon et al. (2000), Lin et al.
128 (2008), Seo et al. (2009), Kim et al. (2009), Vitart and Molteni (2010), Gottschalck et
129 al. (2010), Kang and Kim (2010), Jia et al. (2010), and Crueger et al. (2013). For the

130 medium range (here, roughly +3 to +16 days lead time), the literature on MJO
131 forecast evaluation is sparser. Some verification statistics were calculated for
132 THORPEX Interactive Grand Global Ensemble (TIGGE; Bougeault et al. 2010) data by
133 Matsueda and Endo (2011).
134 Recently, NOAA scientists created an extensive global ensemble reforecast
135 data set (Hamill et al. 2013) using the version of the NCEP Global Ensemble Forecast
136 System (GEFS) that was operational in 2012-2013. This data set was created to
137 facilitate the diagnosis and statistical correction of systematic forecast errors in
138 medium-range ensemble forecasts, thereby improving GEFS guidance. In this article,
139 we demonstrate an ancillary purpose, showing how the extensive reforecasts
140 facilitate the diagnosis of errors in low-frequency modes of variability. Specifically,
141 we will examine the usefulness of the reforecasts for examining the predictability
142 and forecastability of blocking and the MJO as well as their inter-relationships. The
143 operating hypothesis, which as we will show is easily disproved, is that the
144 ensemble prediction system well represents the evolution of forecast uncertainty of
145 these phenomena.
146 Below, section 2 briefly describes the data set and the methods for forecast
147 evaluation. Section 3 provides results, while section 4 provides a discussion and
148 conclusions.
149
150 2. **Data and Methods**
151 *a. Description of the data sets.*

Unless noted otherwise, reforecast data for Dec-Jan-Feb 1985-2012 period was used in this study. The global ensemble reforecast data was more completely described in Hamill et al. (2013). Briefly, this reforecast data set is based on the 2012 version of the NCEP Global Ensemble Forecast System (GEFS). An 11-member retrospective ensemble forecast was generated to +16 days lead for every day at 00 UTC from 1 December 1984 to the current date. Consistent with the operational GEFS, the model resolution was T254L42 to day +8 (~40-km grid spacing at 40° latitude and 42 levels). Starting at day +7.5 and extending to day +16, the reforecasts were conducted at the reduced resolution of T190L42 (~54-km grid spacing). Through 20 February 2011, the Climate Forecast System Reanalysis (CFSR; Saha et al. 2010) provided the control initialization and verification. Thereafter, the operational grid-point statistical interpolation (GSI; Kleist et al. 2009) procedure was used, which was updated to a hybrid variational-ensemble data assimilation approach (Hamill et al. 2011) on 22 May 2012. Additional ensemble member perturbed initial conditions were generated using the ensemble transform with rescaling approach of Wei et al. (2008). See Hamill et al. (2013) for more details on the data set, including a description of the extensive amount of data that is available for fast-access download. In this study, data interpolated to a 1-degree grid was used.

For examination of associated tropical precipitation forecasts, 1-degree Global Precipitation Climatology Project (GPCP, Huffman et al. 2001) data was used. This GPCP data was available only from 1997 – current.

b. Blocking and MJO definitions.

For blocking, though there have been some modern alternatives (e.g., Pelly and Hoskins 2003b, Barnes et al. 2012), here blocking was defined directly following the simple method of Tibaldi and Molteni (1990) based on 500 hPa geopotential heights. Conditional climatologies of blocked and unblocked 500 hPa height patterns at the international dateline are shown in Fig. 2. Given the increased frequency of blocking in the Euro-Atlantic and Pacific regions (Fig. 2a), our analysis of blocking forecastability and predictability was limited to two sectors, the "Euro-Atlantic" sector, from 45° W longitude to 45° E longitude, and the "Pacific" sector, from 140° E longitude to 130° W longitude.

For MJO analysis, a now-standard method was used to compute the projections onto the two leading empirical orthogonal functions (EOFs) of MJO variability (Wheeler and Hendon 2004; hereafter WH04). The respective EOFs are commonly known as "RMM1" and "RMM2." The EOF structures associated with RMM1 and RMM2 were taken directly from the real-time MJO web site, http://cawcr.gov.au/staff/mwheeler/maproom/RMM/ . The EOFs were computed from a time series of longitudinal arrays of filtered anomalies in the 200 and 850 hPa zonal winds and outgoing longwave radiation (OLR), which were averaged in the band from 15° S to 15° N latitude and normalized by their variances over this latitude band and all longitudes. This filtering also removed inter-annual variability and the projection onto the El Niño/Southern Oscillation (ENSO). The processing of the reforecast data to calculate the projections of forecast data onto RMM1 and RMM2 generally followed the procedure outlined by WH04, with the following

198 exception. To remove the effects of ENSO, the projection of the filtered data onto the
199 Niño 3.4 index anomalies was removed (Trenberth 1997), as opposed to the "SST1"
200 index of Drosdowsky and Chambers (2001) cited in WH04. The filtering did
201 remove the leading three harmonics of the annual cycle and the mean of the
202 previous 120 days, as in WH04. Our procedure, when applied to the control initial
203 condition from the CFSR, produced a time series of projections that was a very close
204 match to the time series produced by WH04 with NCEP-NCAR reanalyses (Kalnay et
205 al. 1996) data (not shown).
206
207 c. *Methods of forecast evaluation*.
208 Many methods were used to evaluate blocking and MJO forecast skill. For
209 blocking, a Brier Skill Score was used that measured the skill of the ensemble
210 forecast's ability to set forecast probability of blocked conditions. The Brier Score of
211 the forecast and climatology were computed separately for each longitude in the
212 standard manner (Wilks 2006, eq. 7.34). The Brier Skill Score (BSS; ibid, eq. 7.35)
213 was computed from the sums of Brier Scores of the forecast and climatology over
214 each longitude within a sector. Forecast probabilities of blocking were set directly
215 by ensemble relative frequency. For example, if 3 of the 11 members were
216 diagnosed as having blocked conditions at a particular longitude, the forecast
217 probability was set to 3/11. Since climatological blocking frequency did not vary
218 radically within each sector (see Fig. 2 below), there was little reason to apply more
219 complex procedures (e.g., Hamill and Juras 2006) to account for variations in
220 climatological event frequency for each sector when calculating the BSS. Skill will

221 also be evaluated for block onset and cessation for long-lived blocks. The subset of
222 dates classified as "block onset" for long-lived blocks were those dates where,
223 within a given sector, there were < 20 degrees of longitude that were blocked the
224 day before but ≥ 20 degrees blocked on that day, and where ≥ 20 degrees continued
225 to be blocked for at least the subsequent 9 days. Dates classified as "block cessation"
226 were those dates where, for the first date after a defined block onset, < 20 degrees
227 of longitude were blocked. Finally, we will also evaluate the BSS of a "perfect-model"
228 forecast (Buizza 1997). One of the 11 forecast members replaced the analyzed state,
229 and then the existence / non-existence of a block was computed from this
230 replacement data. Probabilities were estimated from the diagnoses of blocking
231 from the remaining 10 forecast members.

232 Reliability diagrams (Wilks 2006, p. 287) for blocking probabilities were also
233 calculated in the standard manner.

234 The MJO was also evaluated with a variety of metrics. Following now
235 standard metrics defined by Lin et al. (2008), the correlation skill (COR) and root-
236 mean-square error (RMSE) of the RMMs were calculated over all ensemble
237 members as:

238 $$COR(\tau) = \frac{\sum_{j=1}^{11}\sum_{i=1}^{N}\left[RMM1_i^a RMM1_{ij}^f + RMM2_i^a RMM2_{ij}^f\right]}{\sqrt{\sum_{j=1}^{11}\sum_{i=1}^{N}\left[RMM1_i^{a\,2} + RMM2_i^{a\,2}\right]}\sqrt{\sum_{j=1}^{11}\sum_{i=1}^{N}\left[RMM1_{ij}^{f\,2} + RMM2_{ij}^{f\,2}\right]}} \qquad (1)$$

239 and

240 $$\mathrm{RMSE}(\tau) = \sqrt{\frac{1}{\mathrm{Nx11}}\sum_{\mathrm{j=1}}^{11}\sum_{\mathrm{i=1}}^{\mathrm{N}}\left\{\left[RMM1_{\mathrm{i}}^{\mathrm{a}} - RMM1_{\mathrm{ij}}^{\mathrm{f}}\right]^2 + \left[RMM2_{\mathrm{i}}^{\mathrm{a}} - RMM2_{\mathrm{ij}}^{\mathrm{f}}\right]^2\right\}} \qquad (2)$$

where τ is the lead time in days, $RMM1_i^a$ and $RMM2_i^a$ are the analyzed RMM1 and RMM2 projections for the *i*th of *N* sample days. $RMM1_{ij}^f$ and $RMM2_{ij}^f$ are their respective, time-concurrent forecast projections for each of the *j*th of 11 members. The *COR* and *RMSE* will also be calculated for "high-amplitude" and "low-amplitude" events. The *i*th forecast sample is evaluated as high amplitude or low amplitude depending on the magnitude of the associated analysis. The sample is high amplitude if $\sqrt{[RMM1_i^a + RMM2_i^a]^2} \geq 1.0$ and low amplitude if < 1.0.

We were also interested in the ability of the ensemble to provide high-quality, reliable probabilistic guidance. For the MJO, reliability (or more accurately, "consistency") was evaluated with rank histograms (Hamill 2001). Rank histograms were calculated separately for RMM1 and RMM2, but the average of the two is reported. Rank histograms were also generated for MJO phase and amplitude. In generating the rank histograms, the observed value of RMM1 and RMM2 were assumed to be perfect, so no noise was introduced to ensemble members to potentially account for inaccuracies in the analyses of RMMs (ibid, Fig. 6 therein). The phase and amplitude propagation characteristics of the reforecasts will also be examined; the specific methodology for these will be discussed at the relevant point in the results section.

A continuous ranked probability skill score (CRPSS; Wilks 2006, p. 302) was also calculated for MJO forecasts against two reference forecasts, climatology and lagged persistence. As with rank histograms, continuous ranked probability score (CRPS) statistics for the forecast and the reference were calculated separately for RMM1 and RMM2 and then summed. Let Φ_i^f represent the ensemble forecast

cumulative distribution function (CDF) for RMM1 or RMM2 for the ith of N samples. Similarly, let Φ_i^a represent the analyzed CDF, which is a step function, 0 below the RMM value and 1 above. Φ_i^r is the CDF of the reference forecast, be it the unconditional climatology for Dec-Jan-Feb or a lagged persistence forecast. Then the CRPSS was calculated as

$$CRPSS = 1.0 - \frac{CRPS^f}{CRPS^r} \quad , \tag{3}$$

where the $CRPS^f$ was the continuous ranked probability score (CRPS) of the forecast and $CRPS^r$ was the CRPS of the reference. $CRPS^f$ was calculated via

$$CRPS^f = \frac{1}{\mathrm{N}} \sum_{i=1}^{N} \int_{RMM=-\infty}^{\infty} \left(\Phi_i^f - \Phi_i^a\right)^2 dRMM \quad , \tag{4}$$

with the $CRPS^r$ calculated similarly. For the climatological reference forecast, a Gaussian distribution was fitted to the climatology individually for RMM1 and RMM2, which was then used to generate the CDFs. The lagged persistence was based directly on the "PCRLAG" approach documented in Seo et al. (2009, their eq. 1). Future RMM values were predicted using a multiple linear regression model using the current RMM values and their values 5, 10, 15, 20 and 25 days in the past.

Some figures below will include PDFs of the daily change in the angle θ and magnitude of RMM in the (RMM1, RMM2) phase space. θ measures the rotation from the positive RMM1 axis, and is defined as

$$\theta = tan^{-1}(RMM2/RMM1), \tag{5}$$

which can be defined uniquely in the interval $-180° \leq \theta < 180°$ with knowledge of the signs of RMM2 and RMM1. Angles of -180° to -135° correspond to "phase 1" in the (RMM1, RMM2) phase space (see Wheeler and Hendon 2004, Fig. 7); angles of -

135° to -90° correspond to "phase 2," and so on. The PDFs of θ and RMM magnitude were estimated with kernel density estimation using a Cressman-shaped kernel (Cressman 1959) that tapered to 0.0 at 2° or 0.1 units of RMM magnitude.

3. **Results**.

a. Blocking forecasts.

Figure 2a shows the blocking frequency in the GEFS reforecasts for selected forecast lead times. Overall, the GEFS replicated blocking frequency reasonably accurately, though for lead times of +6 days and beyond it under-forecasted blocking frequency in the Euro-Atlantic sector by up to 25%. While the blocking frequency curves are relatively smooth over the multi-decadal period, this disguises tremendous inter-annual variability. Figure 2b shows the yearly blocking frequencies, spatially smoothed slightly to aid in interpretability. For a given longitude, blocking frequencies can vary by an order of magnitude or more from one year to the next. Figure 3 shows that overall positive blocking forecast skill was retained through day +13, but this skill was far short of the skill that was possible under-perfect model assumptions. For example, the perfect-model skill at day +7 was as large as the actual skill at ~ day +3.5. It is somewhat likely that the perfect-model estimate of forecast skill is somewhat too large, too, due to the tendency for the ensemble forecasts to cluster together, such that their spread is not statistically consistent with their ensemble-mean error (e.g., Bougeault et al. 2010, Figs. 2-3). Figure 3 also shows that blocking onset and cessation were somewhat less well forecast than the overall forecasts of blocking. The skill curves for onset and

309 cessation were noisier because of the greatly reduced sample size, even with 28
310 winter seasons of data. Figure 4 shows that there was a substantial amount of
311 variability in the skill of blocking when the data was sorted by half-decadal periods,
312 plus 2010-2012. In the Pacific sector, the blocking skill in the most recent 3 years
313 was the largest, but the 2005-2009 period was intermediate in skill and actually
314 comparable to the skill during the 1985-1989 period. However, in the Euro-
315 Atlantic sector, the skill for the 1985-1989 period was substantially smaller than for
316 the subsequent half decades. The dashed lines in Fig. 4 provide the half-decadal skill
317 results under the perfect-model assumption. These show some natural variability in
318 skill at half-decadal timescales. Note that the 1985-1989 period, for example, had
319 the lowest perfect-model predictability in the Euro-Atlantic sector, which probably
320 contributed to its especially low real-model skill, while the 2010-2012 period had
321 among the highest perfect-model skill in the Pacific, indicating that the actual high
322 forecast skill for this period was in part natural variability. The perfect-model
323 results here also suggest that even in the best of circumstances, blocking forecast
324 skill as defined here is limited to approximately two weeks.
325 Figure 5 presents reliability diagrams for the blocking forecasts at various
326 lead times. At the earlier lead times the blocking forecasts are mostly reliable, but
327 the reliability decreases so that by day +15 the forecasts are rather unreliable. As
328 expected, blocking forecast sharpness decreases over time, as seen in the usage
329 frequency histograms. There are probably many reasons for the lack of reliability,
330 including all the usual suspects with unreliable ensemble forecasts; the moderate
331 resolution of the forecast model, the deficiencies in parameterizations, including

sometimes inappropriate deterministic formulations (Palmer 2012), the sub-optimal initialization of both the control and perturbations in the ensemble system, and, as we shall see below, faulty representations of interactions with the MJO.

b. *MJO forecasts*.

Figure 6 shows the evolution from both analyzed and deterministic forecasts where the vector (RMM1, RMM2) of initial conditions was within 0.5 units of (1.5, -1.5), i.e., within the purple circle on the figure. The samples are all various dates from Dec-Jan-Feb 1985-2012. It appears that the collection of forecasts propagate somewhat more regularly than the collection of analyzed states, and perhaps the forecasts lose some amplitude. Figures 7 and 8 attempt to quantify this, providing a PDF of the daily change in overall RMM angle and magnitude, respectively, as well as the change attributable specifically to wind and OLR components of the RMM. The cases used to populate Figs. 7 and 8 were selected in an attempt to isolate situations where there was a real MJO between Africa and the Maritime continent, and where it had robust associated convection. Consequently, only the Dec-Jan-Feb 1985-2012 cases that were included were those that had an RMM amplitude of greater than 1.0, an RMM2 < 0.0, and an OLR contribution to the RMM amplitude of greater than 0.5 were included. Consider the overall phase change in Fig. 7a; after a reasonable mean phase change the first day, for subsequent leads the mean phase change was substantially smaller for the forecast than for the analyzed, with a phase change of 5-10° for the analyzed but 3-6° for the forecast. Forecast MJOs propagated too slowly, on average. The forecast propagation of the OLR component of RMM was

slower than the wind component, though this was true to a lesser extent with analyzed data as well. This slower propagation can also be diagnosed, for example, from the reanalyzed rainfall and wind lag correlations shown in Fig. 3 from Weaver et al. (2011). By inspection, the distributions of phase changes of the forecast PDFs were not dramatically narrower than of the analyzed PDFs; the main deficiency was a biased mean, not a lack of spread. Considering the overall magnitude change in Fig. 8a, both forecast and analyzed exhibit a similar small decrease in magnitude, though the PDF for the forecast appears slightly more narrow and peaked than for the analyzed.

Not only did the forecast MJO precipitation features propagate too slowly, but also the precipitation forecasts exhibited significant bias (Fig. 9). Precipitation amounts were dramatically over-forecast at the early forecast leads. The general pattern of the daily GPCP precipitation amount climatology was reasonably replicated in day +0 to 1 forecast (Fig. 9b), but the average daily forecast precipitation amounts was commonly > 50% too large. By the beginning of the second week of the forecast (Fig. 9c), the over-forecast bias was reduced, but there was less resemblance with the analyzed precipitation pattern. For example, the connection of the South Pacific Convergence Zone (SPCZ; Folland et al. 2002) to the inter-tropical convergence zone was missing, and the forecast SPCZ was unduly zonally oriented, as it often is in climate simulations (Brown et al. 2011). These pattern changes were even more evident at the end of week +2 (Fig. 9d).

An additional deficiency of the probabilistic MJO forecasts was their under-dispersion and/or conditional bias. This can be seen by examining the rank

histograms from the ensemble predictions of MJO RMMs (Fig. 10). All of the rank histograms were U-shaped, which was most pronounced for the short-lead RMM amplitude forecasts. The rank histograms indicate that there was unrealistic consistency of magnitudes among the forecasts; the ensemble prediction system did not adequately simulate the forecast processes that contribute to diversity in MJO magnitudes.

Despite the significant biases, the forecasts still exhibit skill in the first week. Figure 11 presents the CRPSS of the forecasts measured relative to an unconditional climatology and relative to the regression-based lagged persistence model. The lagged persistence model presented a tougher reference standard, so forecasts exhibited less skill in comparison to this. Skill diminished to near 0 by day +11 with respect to lagged persistence and by day +14 with respect to climatology. While lagged persistence represented a tougher reference, we note that the first-generation Climate Forecast System at NCEP (Wang et al. 2005, Saha et al. 2006) produced forecasts that had higher errors and less correlation skill than the lagged persistence at all forecast leads. Thus, the current GEFS provides substantial improvement in the simulation of the MJO relative to the first-generation CFSR.

Consider now the correlation skill and RMSE of MJO forecasts (Fig. 12). Overall correlation skill and RMSE were comparable to those from the more accurate models shown in Matsueda and Endo (2011). There was more correlation skill in the wind components of the RMM than in the OLR component, as shown in Fig 12(a). This likely relates to the better ability of models to maintain and evolve the rotational component of the wind over the area of convection, perhaps due to

improved initial conditions. The correlation skill was much lower for the first two half-decadal periods, 1985-1989 and 1990-1994, than for the subsequent periods, with the exception of 2010-2012, though the RMS error was not larger for the first decade of the forecast (Figs. 12 (c) and (d)). Figures 12 (e) and (f) show that forecasts initialized from analyses with high amplitude of RMM exhibited more skill but also higher RMSE than lower-amplitude forecasts. This was also shown in Lin et al. (2008, their Fig. 13),

c. Interactions between blocking and the MJO.

Finally, we briefly consider the ability of the forecast model to successfully replicate the ability to discern changes in blocking frequency for different phases of strong MJOs. Strong MJOs are defined as the set of dates where the magnitude of the RMM is in the upper quartile of its distribution, forecast or observed. Figure 13 presents the results. Consider panel (a). Here the change in blocking frequency for a strong MJO relative to the unconditional blocking frequency is shown as a function of longitude (abscissa) and of the analyzed phase θ (ordinate) of the MJO, as defined in section 2c. To provide an adequate sample size for a given θ, the data plotted for a given θ actually includes analyzed samples with similarly diagnosed θ, specifically where $-22.5° \le \theta \le 22.5°$. Note some interesting characteristics in the analyzed relationship of blocking frequency changes. As θ varies between 0° and 120° (i.e., the MJO's center moves from the Maritime continent to western N. America), at 0° longitude, the blocking frequency changes from a strongly negative anomaly in blocking frequency to a strongly positive anomaly. Restated, analyzed Euro-Atlantic

424 blocking frequency changes from below its long-term average to above average as
425 the MJO moves east from the Maritime continent. We note that this is consistent
426 with previous results, such as the 500 hPa anomaly composite for various phases of
427 the MJO in Lin et al. (2009, Fig. 4). Our figure 13(b)-(d) then show the respective
428 blocking frequency anomalies when the *forecast* MJO phase is of the noted angle,
429 and when the *forecast* MJO magnitude is greater than the upper quartile of the
430 forecast distribution. The day +4 forecast in panel (b) still replicates many of the
431 essential anomalies of the analyzed, including the shift from a strongly positive
432 blocking frequency anomaly to a strongly negative anomaly along the Greenwich
433 meridian as θ varies between 0 and 120 degrees. However, much of the frequency
434 anomaly detail is lost by day +8, and the day +16 forecasts show no apparent
435 relation to the analyzed. From this, we can conclude that the internal dynamics of
436 the GEFS do not represent very well the processes that lead to inter-relationships
437 between blocking and the MJO.
438
439 4. **Discussion and conclusions**
440 Some modes of atmospheric variability are uncommon enough and/or
441 operate on long-enough timescales that a short time series of past forecasts will not
442 prove sufficient for diagnosing their characteristics. Atmospheric blocking and the
443 MJO are two such phenomena. In this paper we have shown how a very long time
444 series of ensemble forecast guidance facilitates a greater understanding of the
445 forecastability and predictability of these phenomena. In this case, the long time
446 series was provided by a 28-year data set of reforecasts from the NCEP Global

447 Ensemble Forecast System. The paper more specifically explored Northern
448 Hemispheric blocking, the MJO, and their interaction during Dec-Jan-Feb 1985-2012
449 period.

450 With regards to blocking, the reforecasts showed that the GEFS slightly
451 under-forecasted blocking frequency at longer leads in the Euro-Atlantic sector.
452 Furthermore, the inter-annual variability of blocking frequency was shown to be
453 quite large, demonstrating how difficult it can be to achieve a representative sample
454 with only a few years of data. The predictive skill of the probabilistic forecasts of
455 actual blocking was substantially smaller than its perfect-model skill, whereby a
456 member of the ensemble was used as a synthetic verification. This indicates that
457 there is still tremendous potential for improvement in blocking forecasts. However,
458 it is also likely that the perfect-model results present a somewhat over-optimistic
459 estimate of the upper range of forecast skill. The GEFS system and most other
460 ensemble systems are under-dispersive, and as such, the members of the ensemble
461 unduly resemble each other, inflating the perfect-model skill estimates. It was also
462 found that block onset and cessation were forecast somewhat less well than block
463 maintenance, and there was substantial variability of blocking skill between half-
464 decadal periods. Finally, the reliability of probabilistic blocking forecasts degraded
465 with increasing lead time, and as expected, blocking forecasts became progressively
466 less sharp, i.e., forecast probabilities were less often 0.0 and 1.0 and more often
467 resembled the model climatology.

468 Forecasts of strong MJOs propagated too slowly, especially the component
469 associated with outgoing longwave radiation (OLR), i.e., convection. Deep tropical

convection appeared to have other systematic biases in the GEFS; in general, there was too much tropical precipitation forecast in the Indian and Pacific Oceans, especially for the shorter forecast leads. The ensemble predictions were biased and/or under-dispersive, manifested in U-shaped rank histograms of MJO indices. Forecasts of the magnitude of the MJO's leading EOFs were especially U-shaped. Bi-variate MJO correlation skill was found to be larger for the wind component than for OLR component, and skill was larger for the higher-amplitude MJO events. Skill varied significantly between half-decadal periods, with the period 1985-1994 and 2010-2012 exhibiting lower MJO skill than the 1995-2009 period. Probabilistic skill of the MJO forecast was modest, and skill was larger when measured relative to climatology than when measured relative to a lagged persistence forecast. Finally, for longer-lead forecasts, the GEFS demonstrated little ability to replicate the changes in blocking frequency due to a strong MJO that were noted in analyzed data.

This paper has discussed forecast skill without providing analysis of the potential forecast systematic errors that may lead to deficiencies in blocking, the MJO, and their inter-relationships. This much more challenging work is left as future research. We do hope that we have laid out a first step, demonstrating the predictability and forecastability of these phenomena using the newly created GEFS reforecast data set.

Acknowledgments:

The reforecast data set used here was created under a 2010 Department of Energy Advanced Leadership Computing Challenge supercomputer grant. This manuscript

493 was partly supported by a NOAA USWRP grant on methods for estimating model
494 uncertainty. We thank John Cortinas, the NOAA program manager, for his assistance
495 and guidance on USWRP.
496
497

References:

Barnes, E., A., J. Slingo, and T. Woolings, 2012: A methodology for the comparison of blocking climatologies across indices, models, and climate scenarios. *Clim. Dyn*., **38**, 2467-2481. DOI: 10.1007/s00382-011-1243-6.

Bougeault, P., and others, 2010: The THORPEX interactive grand global ensemble. *Bull. Amer. Meteor. Soc*., **91**, 1059-1072, doi: 10.1175/2010BAMS2853.1.

Brown, J. R., S. B. Power, F. P. Delage, R. A. Colman, A. F. Moise, and B. F. Murphy, 2011. Evaluation of the South Pacific Convergence Zone in IPCC AR4 climate model simulations of the Twentieth Century. *J. Climate*, **24**, 1565-1582.

Buizza, R., 1997: Potential forecast skill of ensemble prediction and spread and skill distributions of the ECMWF ensemble prediction system. *Mon. Wea. Rev*., **125**, 99-119.

Crueger, T. , B. Stevens, R. Brokopf, 2013: The Madden-Julian Oscillation in ECHAM6 and the introduction of an objective MJO metric. *J. Climate*, in press. Available at http://journals.ametsoc.org/doi/pdf/10.1175/JCLI-D-12-00413.1

Drosdowsky, W., and L. E. Chambers, 2001: Near-global sea surface temperature anomalies as predictors of Australian seasonal rainfall. *J. Climate*, **14**, 1677-1687.

Ferranti, L., Palmer, T. N., Molteni, F., and Klinker, E., 1990: Tropical-extratropical interaction associated with the 30-60 day oscillation and its impact on medium and extended-range prediction. *J. Atmos. Sci*, **125**, 2177-2199.

520 Folland, C. K., J. A. Renwick, M. J. Salinger, and A. B. Mullan, 2002: Relative influences
521 of the interdecadal Pacific Oscillation and ENSO in the South Pacific
522 Convergence Zone. *Geophys. Res. Letters*, **29**, 21-1-21-4. Doi:
523 10.1029/2011GL014201.
524 Gottschalck, J., M. Wheeler, K. Weickmann, F. Vitart, N. Savage, H. Lin, H. Hendon, D.
525 Waliser, K. Sperber, M. Nakagawa, C. Prestrelo, M. Flatau, and W. Higgins,
526 2010: A framework for assessing operational Madden-Julian Oscillation
527 forecasts. *Bull. Amer. Meteor. Soc*., **91**, 1247-1258.
528 Hamill, T. M., 2001: Interpretation of rank histograms for verifying ensemble
529 forecasts. *Mon. Wea. Rev*., **129**, 550-560.
530 Hamill, T. M., and J. Juras, 2006: Measuring forecast skill: is it real skill or is it the
531 varying climatology? *Quart. J. Royal Meteor. Soc*., **132**, 2905-2923.
532 Hamill, T. M., J. S. Whitaker, D. T. Kleist, M. Fiorino, and S. G. Benjamin, 2011:
533 Predictions of 2010's tropical cyclones using the GFS and ensemble-based
534 data assimilation methods. *Mon. Wea. Rev*., **139**, 3243-3247.
535 Hamill, T. M., G. T. Bates, J. S. Whitaker, D. R. Murray, M. Fiorino, T. J. Galarneau, Jr., Y.
536 Zhu, and W. Lapenta, 2013: NOAA's second-generation global medium-range
537 ensemble reforecast data set. *Bull. Amer. Meteor. Soc*., **94**, in press.
538 DOI:10.1175/BAMS-D-12-00014.1
539 Hendon, H. H., B. Liebmann, M. Newman, J. D. Glick, 2000: Medium-range forecast
540 errors associated with active episodes of the Madden-Julian Oscillation. *Mon.*
541 *Wea. Rev*., **128**, 69-86.

542 Huffman, G.J., R.F. Adler, M. Morrissey, D.T. Bolvin, S. Curtis, R. Joyce, B McGavock, J.
543 Susskind, 2001: Global precipitation at one-degree daily resolution from
544 multi-satellite observations. *J. Hydrometeor*., **2**, 36-50.
545 Jia, X., C. Li, N. Zhou, and J. Ling, 2010: The MJO in an AGCM with three different
546 cumulus parameterization schemes. *Dyn. of Atmos. Oceans*, **49**, 141-163.
547 Jones, C., D. E. Waliser, K. M. Lau, and W. Stern, 2004: The Madden-Julian Oscillation
548 and its impact on Northern Hemisphere weather predictability. *Mon. Wea.*
549 *Rev*., **132**, 1462-1471.
550 Jung, T., M. J. Miller, T. N. Palmer, P. Towers, N. Wedi, D. Achuthavarier, J. M. Adams,
551 B. A. Cash, J. L. Kinter III, L. Marx, C. Stan, and K. I. Hodges, 2012: High-
552 resolution global climate simulations with the ECMWF model in project
553 Athena: experimental design, model climate, and seasonal forecast skill. *J.*
554 *Climate*, **25**, 3155-3172. DOI: 10.1175/JCLI-D-11-00265.1
555 Kalnay, E., and co-authors, 1996: The NCEP/NCAR 40-year reanalysis project. *Bull.*
556 *Amer. Meteor. Soc*., **77**, 437-471.
557 Kang., I.-S., and H.-M. Kim, 2010: Assessment of MJO predictability for Boreal winter
558 with various statistical and dynamical models. *J. Climate*, **23**, 2368-2378.
559 Kiladis, G. N., and K. M. Weickmann, 1992: Circulation anomalies associated with
560 tropical convection during norther winter. *Mon. Wea. Rev*., **120**, 1900-1923.
561 Kim, D., K. Sperber, W. Stern, D. Waliser, I. S. Kang, E. Maloney, W. Wang, K.
562 Weickmann, J. Benedict, M. Khairoutdinov, M.-I. Lee, R. Neale, M. Suarez, K.
563 Thayer-Calder, and G. Zhang, 2009: Application of MJO simulation
564 diagnostics to climate models. *J. Climate*, **22**, 6413-6436.

565 Kleist, D. T., D. F. Parrish, J. C. Derber, R. Treadon, W.-S. Wu, and S. Lord, 2009:
566 Introduction of the GSI into the NCEP global data assimilation system. *Wea.*
567 *Forecasting*, **24**, 1691-1705.
568 Knutson, T. R. and K. M. Weickmann, 1987: 30-60 day atmospheric oscillations:
569 Composite life cycles of convection and circulation anomalies. *Mon. Wea. Rev*.,
570 **115**, 1407-1436.
571 Leibmann, B., and D. L. Hartmann, 1984: An observational study of tropical-
572 midlatitude interaction on intraseasonal timescales during winter. *J. Atmos.*
573 *Sci.*, **41**, 3333-3350.
574 Lin, H., G. Brunet, and J. Derome, 2008: Forecast skill of the Madden-Julian
575 Oscillation in two Canadian atmospheric models. *Mon. Wea. Rev*., **136**, 4130-
576 4149.
577 Lin, H., G. Brunet, and J. Derome, 2009: An observed connection between the North
578 Atlantic Oscillation and the Madden-Julian Oscillation. *J. Climate*, **22**, 364-380.
579 Madden, R. L., and Julian, P. R., 1971: Detection of a 40-50 day oscillation in the zonal
580 wind in the tropical Pacific. *J. Atmos. Sci*., **5**, 702-708.
581 Maloney, E. D., and Hartmann, D. L., 2000a: Modulation of the eastern North Pacific
582 hurricanes by the Madden-Julian Oscillation. *J. Climate*, **8**, 1757-1774.
583 Maloney, E. D., and Hartmann, D. L., 2000b: Modulation of hurricane activity in the
584 Gulf of Mexico by the Madden-Julian Oscillation. *Science*, **287**, 2002-2004.
585 Matsueda, M., and Endo, H., 2011: Verification of medium-range MJO forecasts with
586 TIGGE. *Geophys. Res. Letters*, **38**, 1-6. Doi:10.1029/2011GL047480.

587 Mauritsen, T., and E. Källén, 2004: Blocking prediction in an ensemble forecasting
588 system. *Tellus*, **56A**, 218-228.
589 Palmer, T. N., 2012: Towards the probabilistic earth-system simulator: a vision for
590 the future of climate and weather prediction. *Quart. J. Royal Meteor. Soc*., **138**,
591 841-861. DOI: 10.1002/qj.1923.
592 Pelly, J. L., and B. J. Hoskins, 2003a: How well does the ECMWF ensemble prediction
593 system predict blocking? *Quart. J. Royal Meteor. Soc*., **129**, 1683-1702.
594 Pelly, J. L., and B. J. Hoskins, 2003b: A new perspective on blocking. *J. Atmos. Sci*., **60**,
595 743-755.
596 Riddle, E. E., M. B. Stoner, N. C. Johnson, M. L. L'Heureux, D. C. Collins, and S. B.
597 Feldstein, 2013: The impact of the MJO on clusters of wintertime circulation
598 anomalies over the North American Region. *Clim. Dyn*., **40**, 1749-1766, DOI:
599 10.1007/s00382-102-1493-y.
600 Saha, S., and co-authors, 2006: The NCEP Climate Forecast System. *J. Climate*, **19**,
601 3483-3517.
602 Saha, S., and co-authors, 2010: The NCEP climate forecast system reanalysis. *Bull.*
603 *Amer. Meteor. Soc.*, **91**, 1015-1057.
604 Seo, K.-H., W. Wang, J. Gottschalck, Q. Zhang, J.-K., Schemm, W. R. Higgins, and A.
605 Kumar, 2009: Evaluation of MJO forecast skill from several statistical and
606 dynamical forecast models. *J. Climate*, **22**, 2372-2388.
607 Straub, K. H., 2013: MJO initiation in the realtime multivariate MJO index. *J. Climate*,
608 **26**, 1130-1151.

609 Tibaldi, S., and F. Molteni, 1990: On the operational predictability of blocking.
610 *Tellus*, **42A**, 343-365.
611 Trenberth, K. E., 1997: The definition of El Niño. *Bull. Amer. Meteor. Soc*., **78**, 2771-
612 2777.
613 Vitart, F., and F. Molteni, 2010: Simulation of the Madden-Julian Oscillation and its
614 teleconnections in the ECMWF forecast system. *Quart. J. Royal Meteor. Soc*.,
615 **136**, 842-855.
616 Waliser, D., and others, 2009: MJO simulation diagnostics. *J. Climate*, **22**, 3006-
617 3030.
618 Wang, W., S. Saha, H.-L. Pan, S. Nadiga, and G. White, 2005: Simulation of ENSO in the
619 new NCEP Coupled Forecast System Model (CFS03). *Mon. Wea. Rev*., **133**,
620 1574-1593.
621 Watson, J. S., and S. J. Colucci, 2002: Evaluation of ensemble predictions of blocking
622 in the NCEP global spectral model. *Mon. Wea. Rev*., **130**, 3008-3021.
623 Weaver, S. J., W. Wang, M. Chen, and A. Kumar, 2011: Representation of MJO
624 variability in the NCEP climate forecast system. *J. Climate*, **24**, 4676-4694.
625 Wei, M., Z. Toth, R. Wobus, and Y. Zhu, 2008: Initial perturbations based on the
626 ensemble transform (ET) technique in the NCEP global operational forecast
627 system. *Tellus A*, **60**, 62-79.
628 Weickmann, K., and E. Berry, 2009: The tropical Madden–Julian Oscillation and the
629 global wind oscillation. *Mon. Wea. Rev.*, **137**, 1601–1614.
630 DOI: http://dx.doi.org/10.1175/2008MWR2686.1

631 Wheeler, M. C., and H. H. Hendon, 2004: An all-season real-time multivariate MJO
632 index: development of an index for monitoring and prediction. *Mon. Wea.*
633 *Rev*., **132**, 1917-1932.
634 Wilks, D. S., 2006: *Statistical Methods in the Atmospheric Sciences (2nd Ed.)*.
635 Academic Press, 627 pp.
636 Yasunari, T., 1979: Cloudiness fluctuations associated with the Northern
637 Hemisphere summer monsoon. *J. Meteor. Soc. Japan*, **58**, 225-229.
638
639

FIGURE CAPTIONS

Figure 1: Composite of Northern Hemisphere 500 hPa geopotential height patterns under (a) blocked flow at 180° E longitude, and (b) unblocked flow at 180° E longitude.

Figure 2: (a) Blocking frequency as determined from analyses and reforecasts as a function of forecast lead time. Areas shaded in gray denote the two sectors in subsequent figures, the Pacific and Euro-Atlantic sectors. (b) Analyzed, spatially smoothed yearly blocking frequencies for each year between 1985 and 2012.

Figure 3: Brier Skill Scores of blocking probability forecasts for (a) Pacific, and (b) Euro-Atlantic sectors.

Figure 4: As in Fig. 3, but for Brier Skill Scores of blocking probability forecasts by half decade for (a) Pacific, and (b) Euro-Atlantic sectors. Solid lines present the skill scores for the actual reforecasts, dashed lines present the skill under perfect-model assumptions.

Figure 5: Reliability diagrams for blocking probability forecasts for (a) +3 day forecast, (b) +6 day forecast, (c) +9 day forecast, (d) +12 day forecast, and (e) +15 day forecast. Dotted red line denotes the skill in the Euro-Atlantic sector and dotted blue line denotes the skill in the Pacific sector. Red and blue bars indicate the frequency of usage of each forecast probability category for the Euro-Atlantic and Pacific sectors, respectively.

Figure 6: RMM1 and RMM2 phase plots for (a) analyzed, and (b) control forecasts whose initial states are within the purple circle. Differences in time from the initial

time conveyed by the colors of lines and dots, with the legend indicating the lead time in days. Quadrants 1 to 8 are marked in the corners of the diagram.

Figure 7: PDF of daily change in angle of RMM vector ($\Delta\theta$), measured in degrees, for (a) overall RMM, (b) OLR contribution to RMM, and (c) wind-component contribution to RMM. Only Dec-Jan-Feb 1985-2012 dates which had an initial RMM magnitude of greater than 1.0, an RMM2 component < 0.0, and an initial OLR component of > 0.5 were included as samples. Dots indicate the mean of the PDF for a given day, horizontally offset slightly so that dots do not overlap.

Figure 8: As in Fig. 7, but for the PDF of the change in magnitude of RMM.

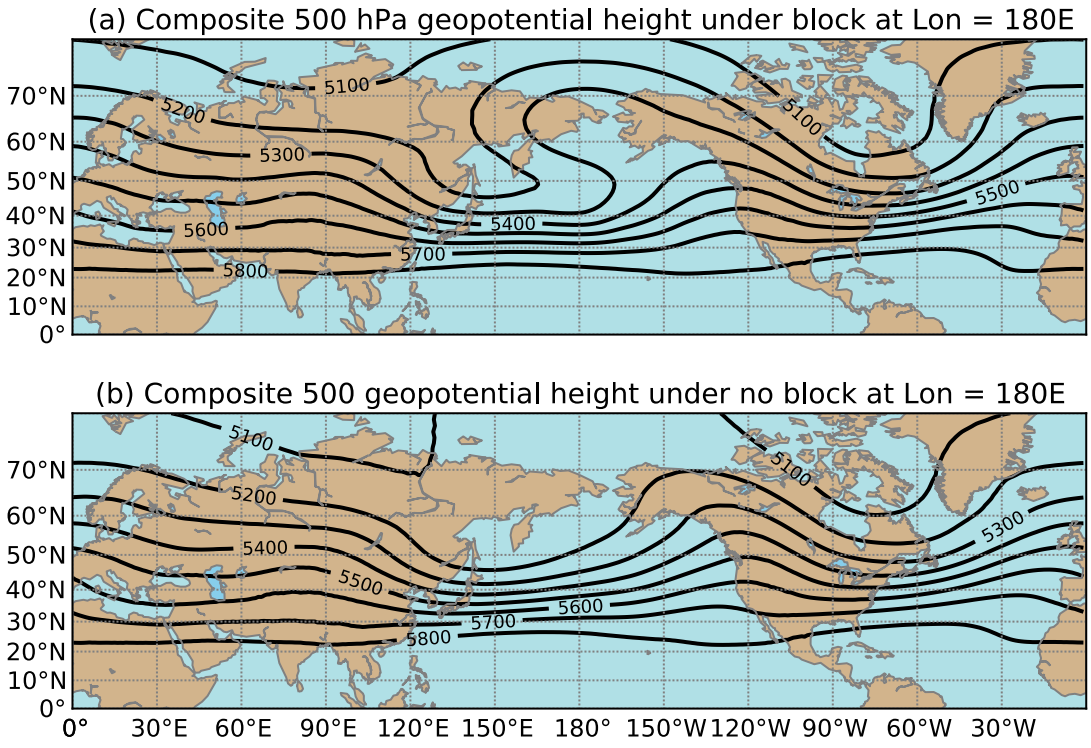
Figure 9: Dec-Jan-Feb 1997-2012 daily-average precipitation climatologies for (a) analyzed GPCP data, (b) +0 to 1 day forecast, (c) +7 to 8-day forecast, and (d) +15 to 16-day forecast.

Figure 10: Rank histograms of RMM1 and RMM2 values [green bars], the angle of the vector in the (RMM1, RMM2) phase space [red bars], and the magnitude [blue bars], for forecast lead times of 1 to 16 days.

Figure 11: Continuous ranked probability skill scores (CRPSS) of ensemble reforecasts of the MJO relative to an unconditional climatological distribution (red line) and relative to a lagged regression model using current and recent analyzed RMM values as predictors (blue line). The green line shows the "perfect-model" skill, when one forecast member is used as a surrogate for the verification and the remaining 10 members are used to generate the probabilities.

Figure 12: Correlation skill (panels (a), (c), and (e)) and RMSE (panels (b), (d), and (f)) for MJO forecasts. Panels (a) and (b) show skill and RMSE for total and for

686 individual wind and OLR components of RMM. Panels (c) and (d) show overall skill
687 and RMSE for half-decadal periods. Panels (e) and (f) show skill for cases with
688 initial large and small amplitude, as defined in the text.
689 **Figure 13**: Change in blocking frequency as a function of longitude for when
690 analyzed (panel a) or +4, +8, or +16 day forecast (panels b, c, and d, respectively)
691 has RMM of phase theta + / - 22.5 degrees. Quadrants of MJO diagram as marked on
692 Fig. 6 are noted on the right-hand side of the plots. The ranges the Euro-Atlantic and
693 Pacific Sectors are noted with heavy black lines at the bottom of each panel.
694
695



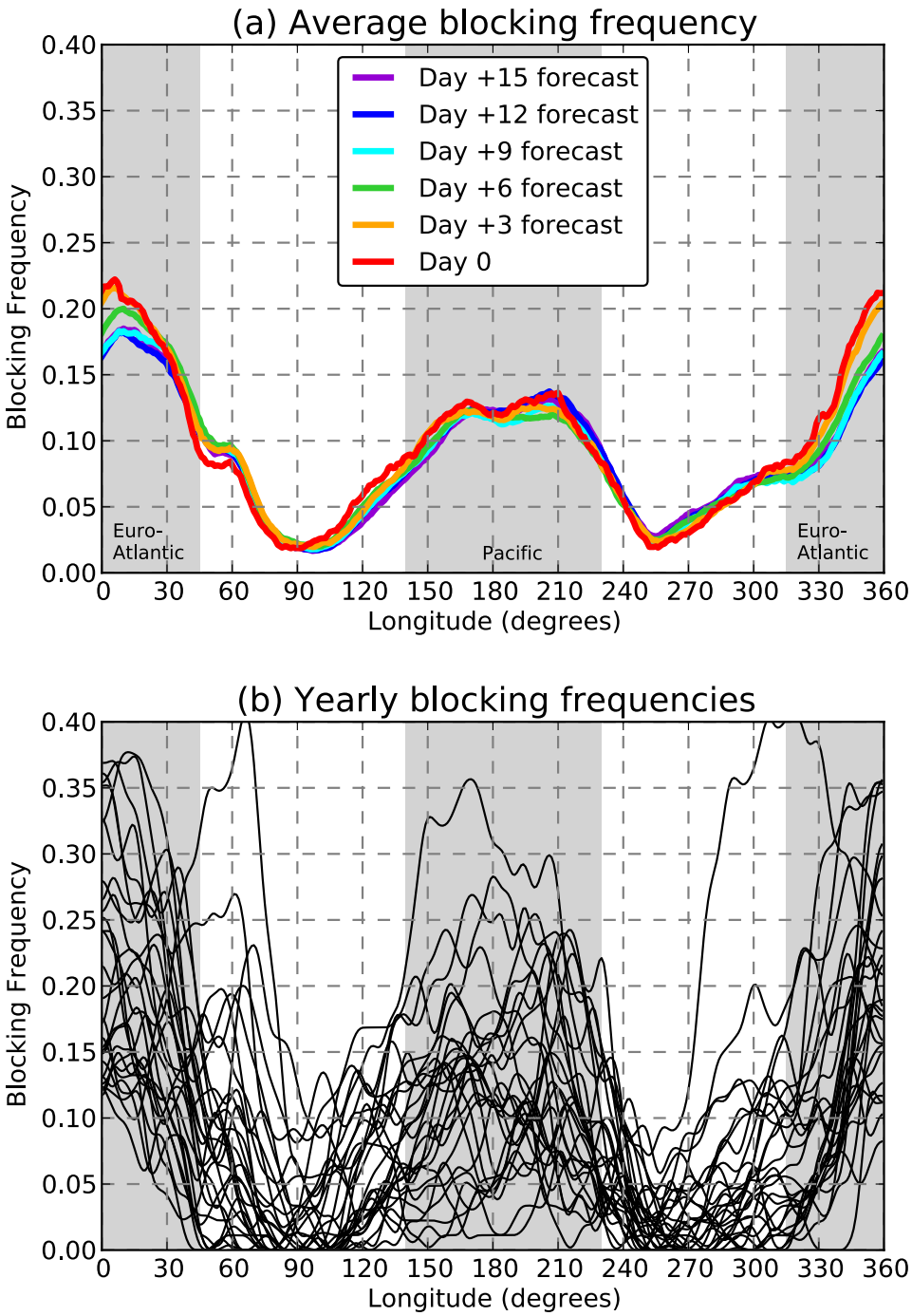
696

697 **Figure 1:** Composite of Northern Hemisphere 500 hPa geopotential height patterns

698 under (a) blocked flow at 180° E longitude, and (b) unblocked flow at 180° E

699 longitude.

700



701

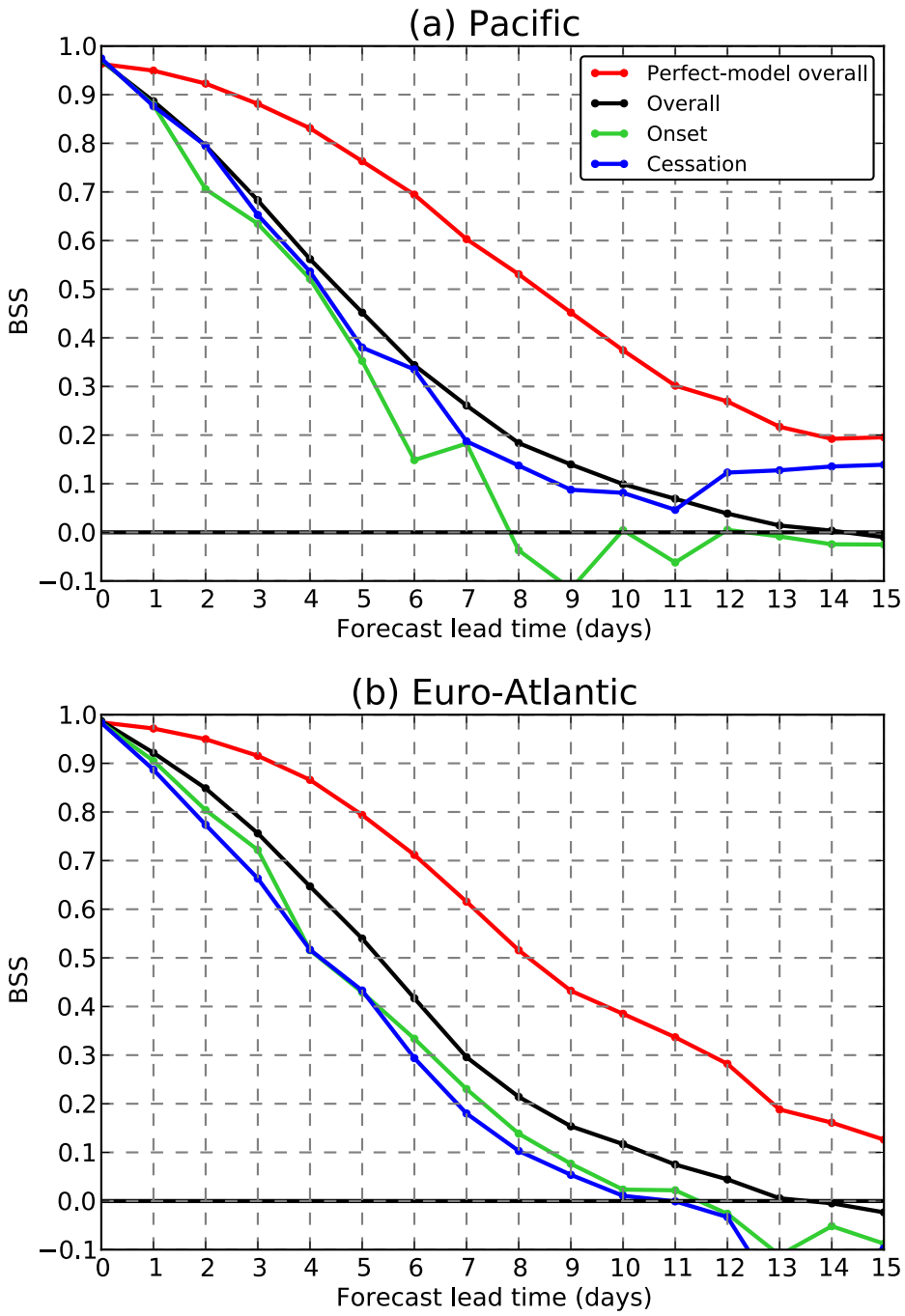
702 **Figure 2:** (a) Blocking frequency as determined from analyses and reforecasts as a

703 function of forecast lead time. Areas shaded in gray denote the two sectors in

704 subsequent figures, the Pacific and Euro-Atlantic sectors. (b) Analyzed, spatially

705 smoothed yearly blocking frequencies for each year between 1985 and 2012.

706

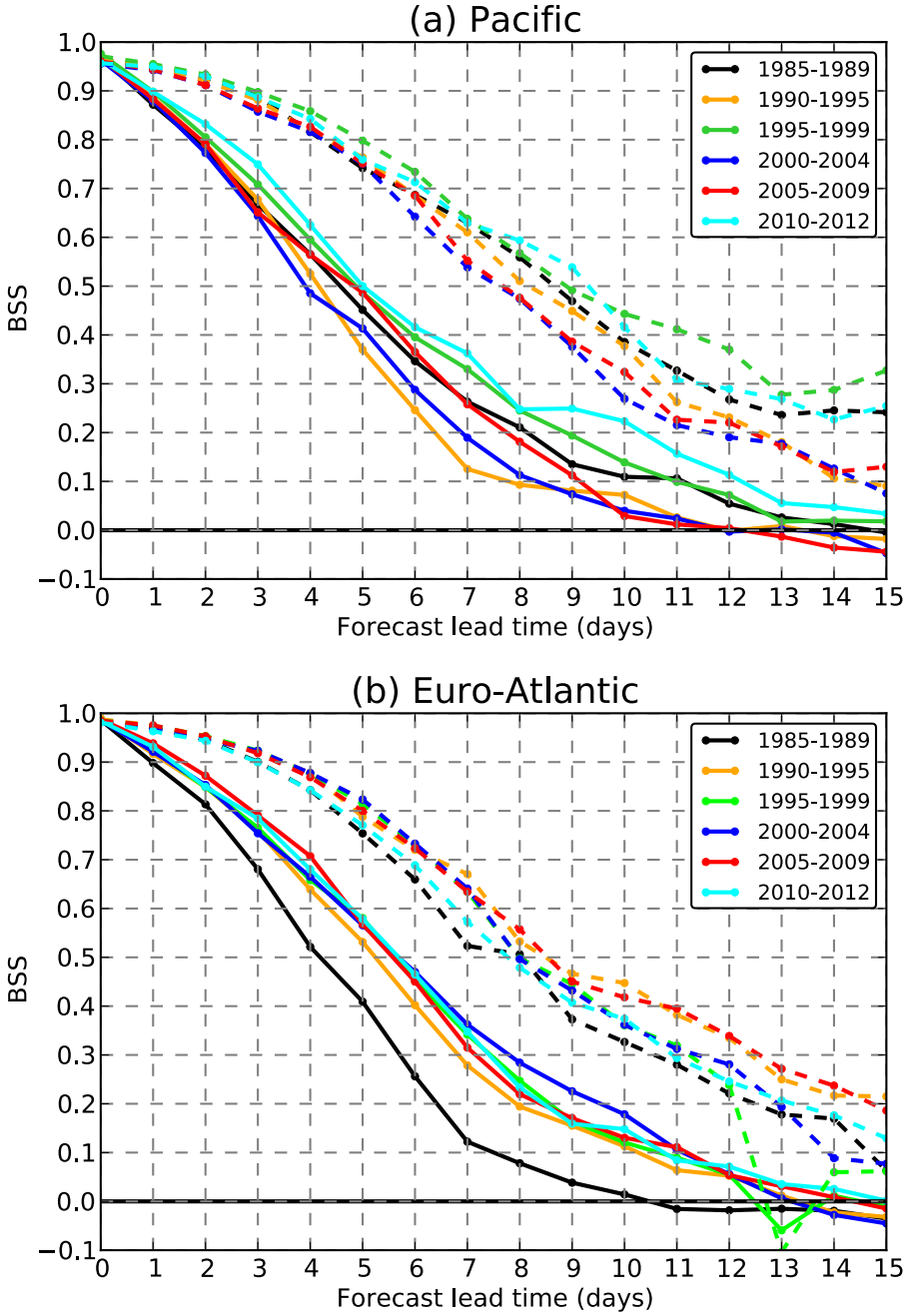


707

708 **Figure 3**: Brier Skill Scores of blocking probability forecasts for (a) Pacific, and (b)

709 Euro-Atlantic sectors.

710



711

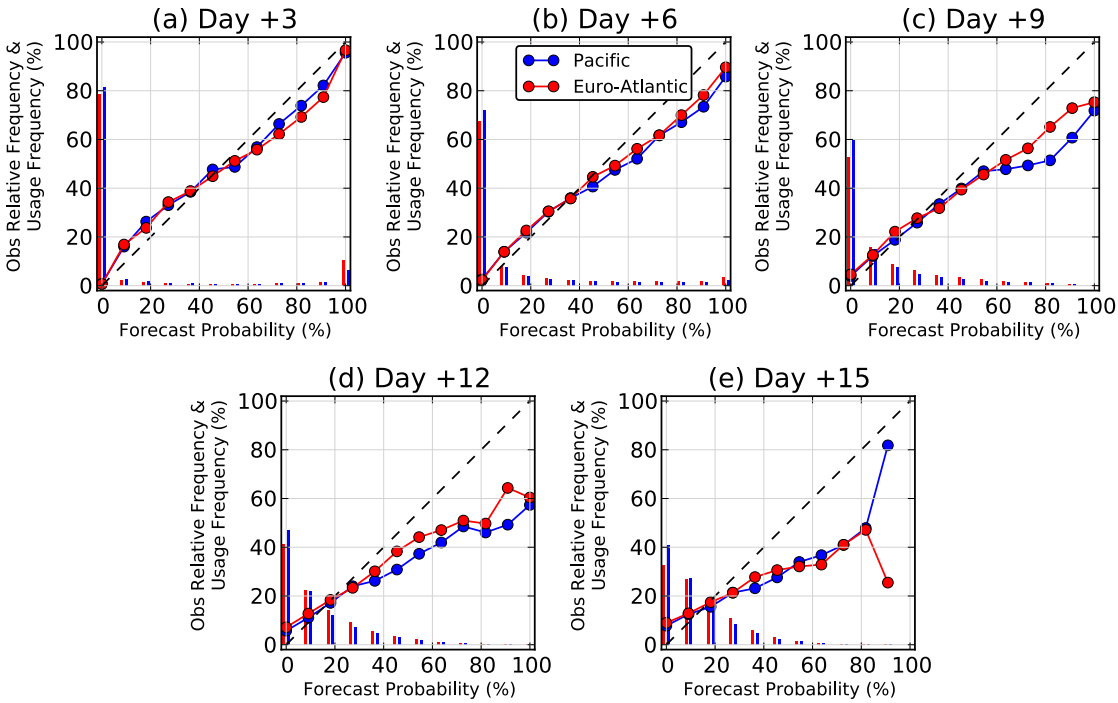
712 **Figure 4**: As in Fig. 3, but for Brier Skill Scores of blocking probability forecasts by

713 half decade for (a) Pacific, and (b) Euro-Atlantic sectors. Solid lines present the skill

714 scores for the actual reforecasts, dashed lines present the skill under perfect-model

715 assumptions.

716



717

718

719 **Figure 5**: Reliability diagrams for blocking probability forecasts for (a) +3 day

720 forecast, (b) +6 day forecast, (c) +9 day forecast, (d) +12 day forecast, and (e) +15

721 day forecast. Dotted red line denotes the skill in the Euro-Atlantic sector and dotted

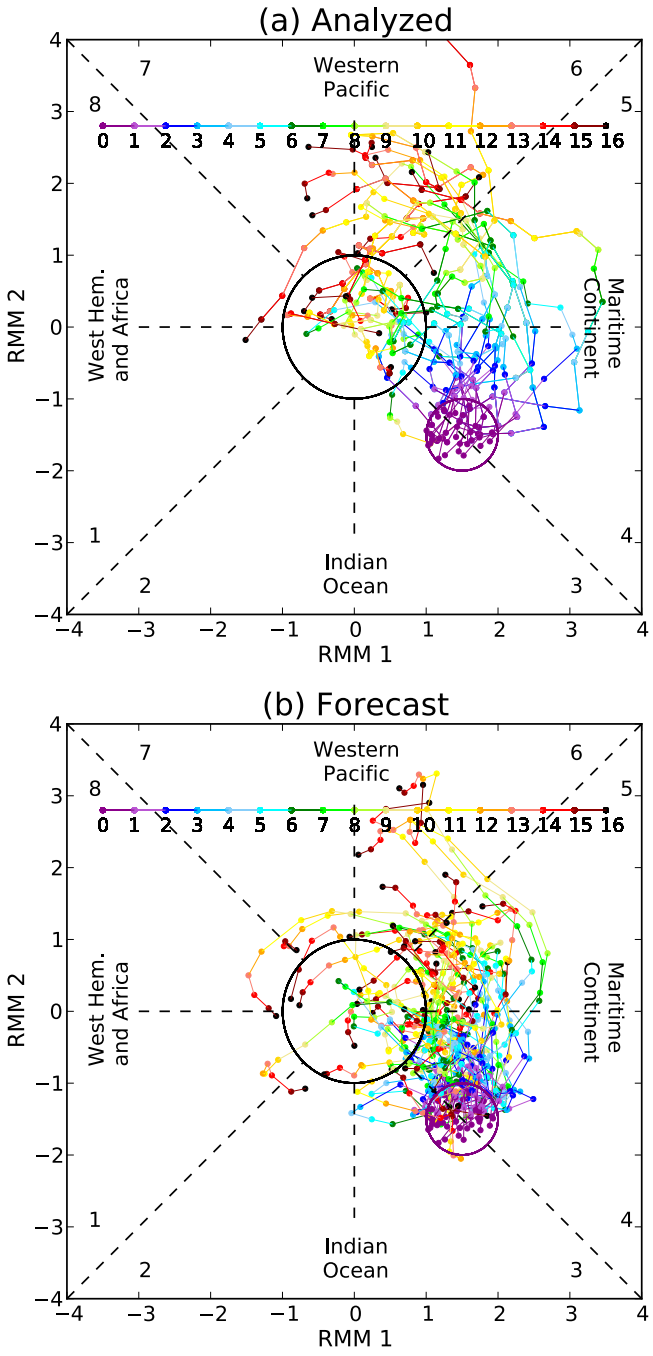
722 blue line denotes the skill in the Pacific sector. Red and blue bars indicate the

723 frequency of usage of each forecast probability category for the Euro-Atlantic and

724 Pacific sectors, respectively.

725

726



727

728 **Figure 6**: RMM1 and RMM2 phase plots for (a) analyzed, and (b) control forecasts

729 whose initial states are within the purple circle. Differences in time from the initial

730 time conveyed by the colors of lines and dots, with the legend indicating the lead

731 time in days. Quadrants 1 to 8 are marked in the corners of the diagram.

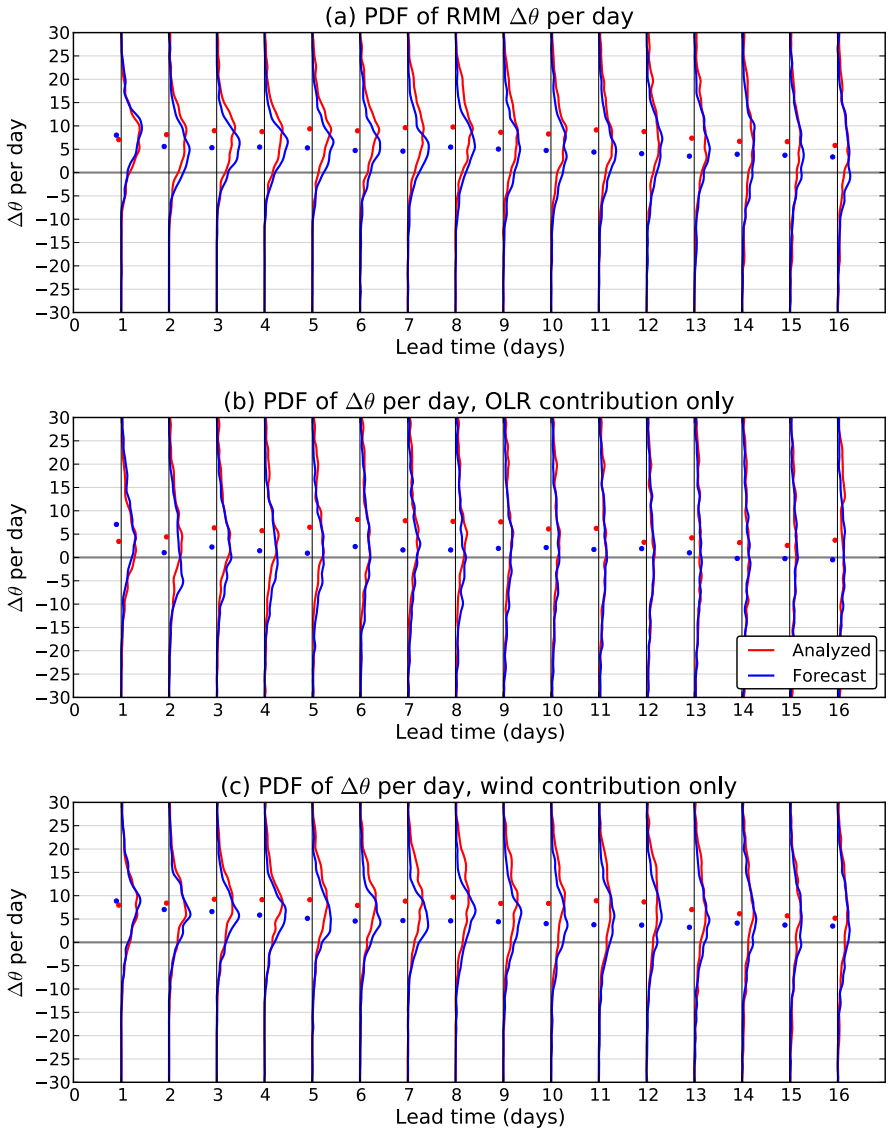
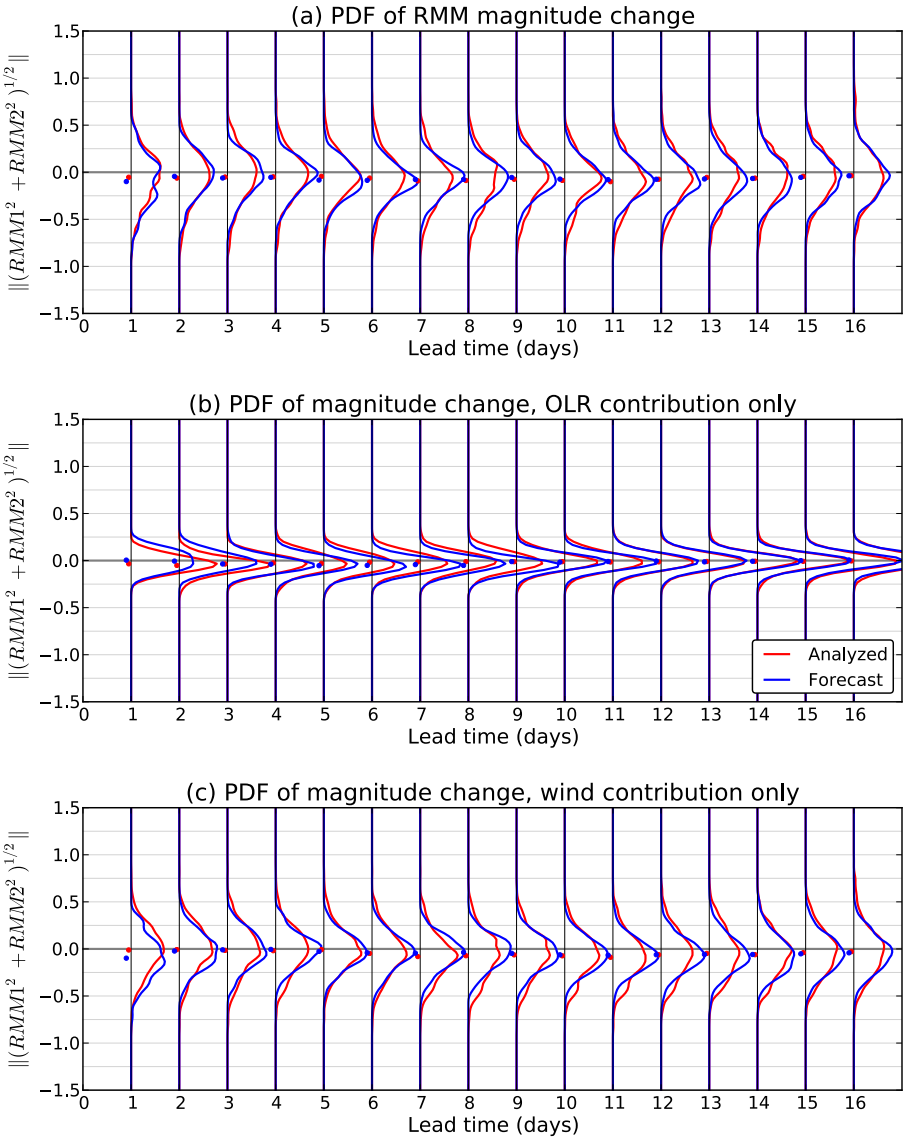


Figure 7: PDF of daily change in angle of RMM vector (Δθ), measured in degrees, for (a) overall RMM, (b) OLR contribution to RMM, and (c) wind-component contribution to RMM. Only Dec-Jan-Feb 1985-2012 dates which had an initial RMM magnitude of greater than 1.0, an RMM2 component < 0.0, and an initial OLR component of > 0.5 were included as samples. Dots indicate the mean of the PDF for a given day, horizontally offset slightly so that dots do not overlap.

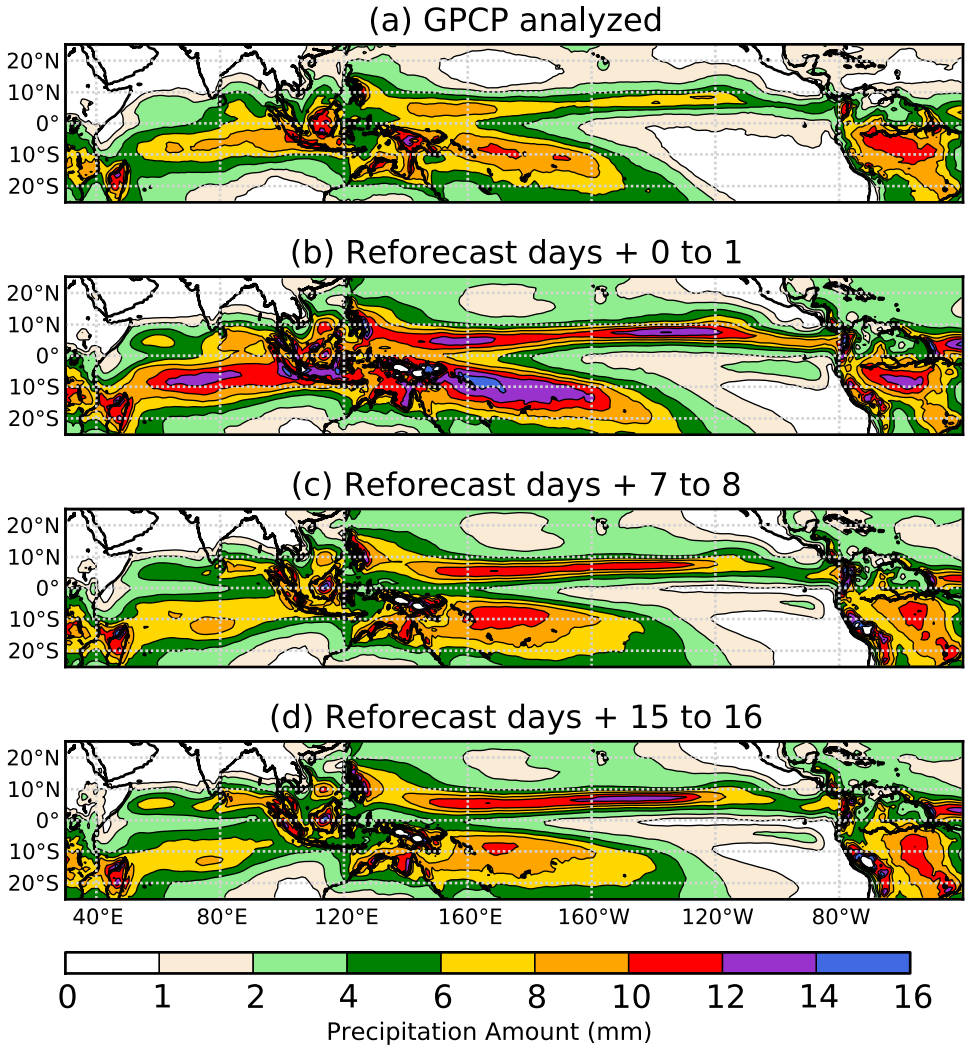
740



741

742 **Figure 8**: As in Fig. 7, but for the PDF of the change in magnitude of RMM.

743



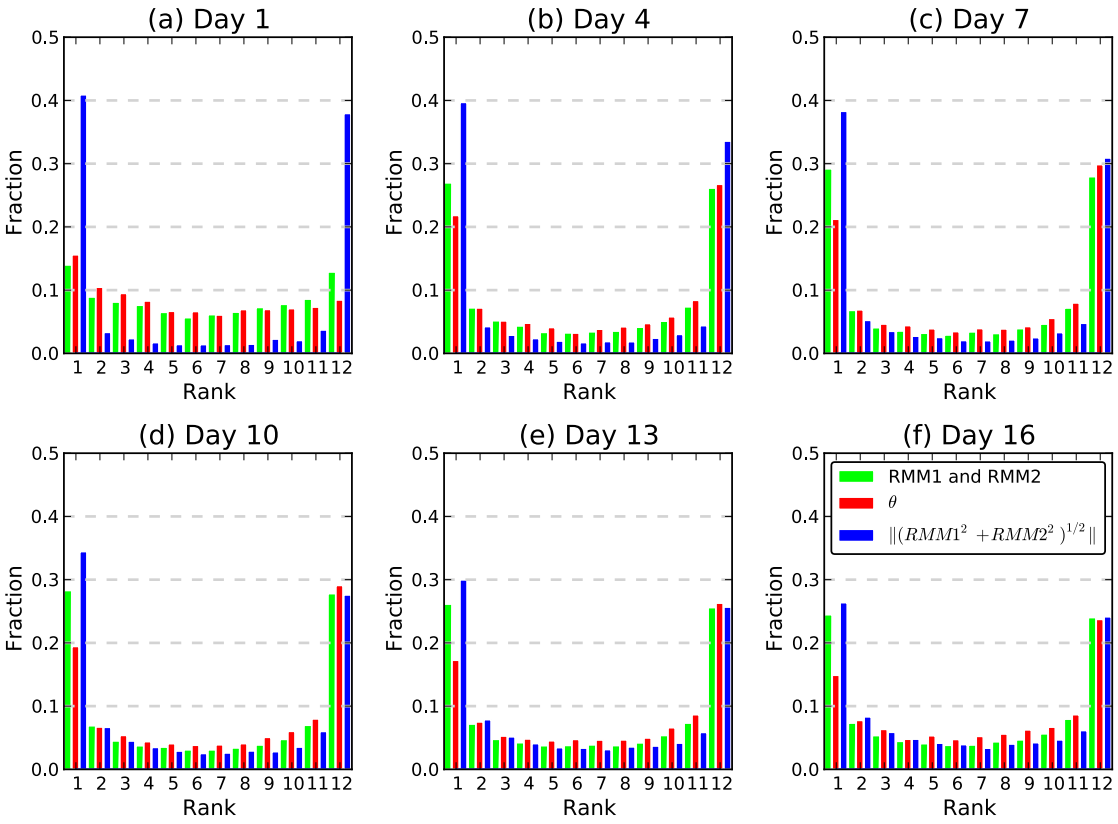
744

745 **Figure 9**: Dec-Jan-Feb 1997-2012 daily-average precipitation climatologies for (a)

746 analyzed GPCP data, (b) +0 to 1 day forecast, (c) +7 to 8-day forecast, and (d) +15 to

747 16-day forecast.

748



749

750 **Figure 10**: Rank histograms of RMM1 and RMM2 values [green bars], the angle of

751 the vector in the (RMM1, RMM2) phase space [red bars], and the magnitude [blue

752 bars], for forecast lead times of 1 to 16 days.

753

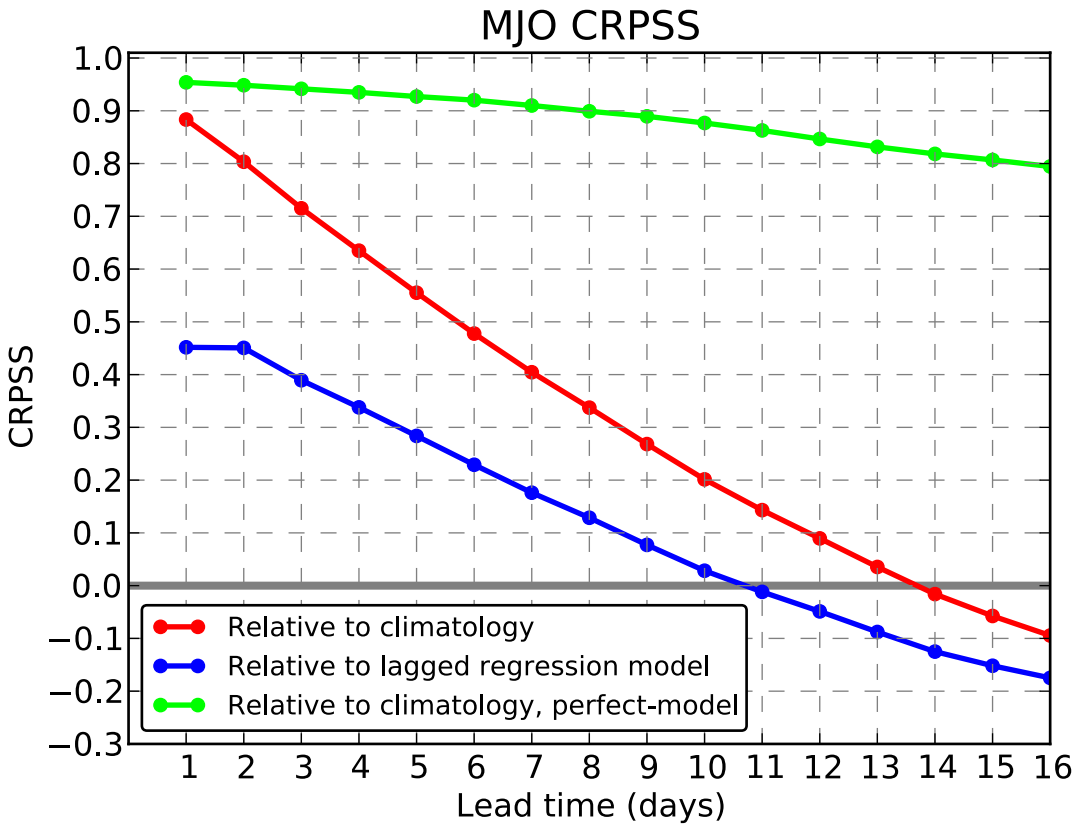
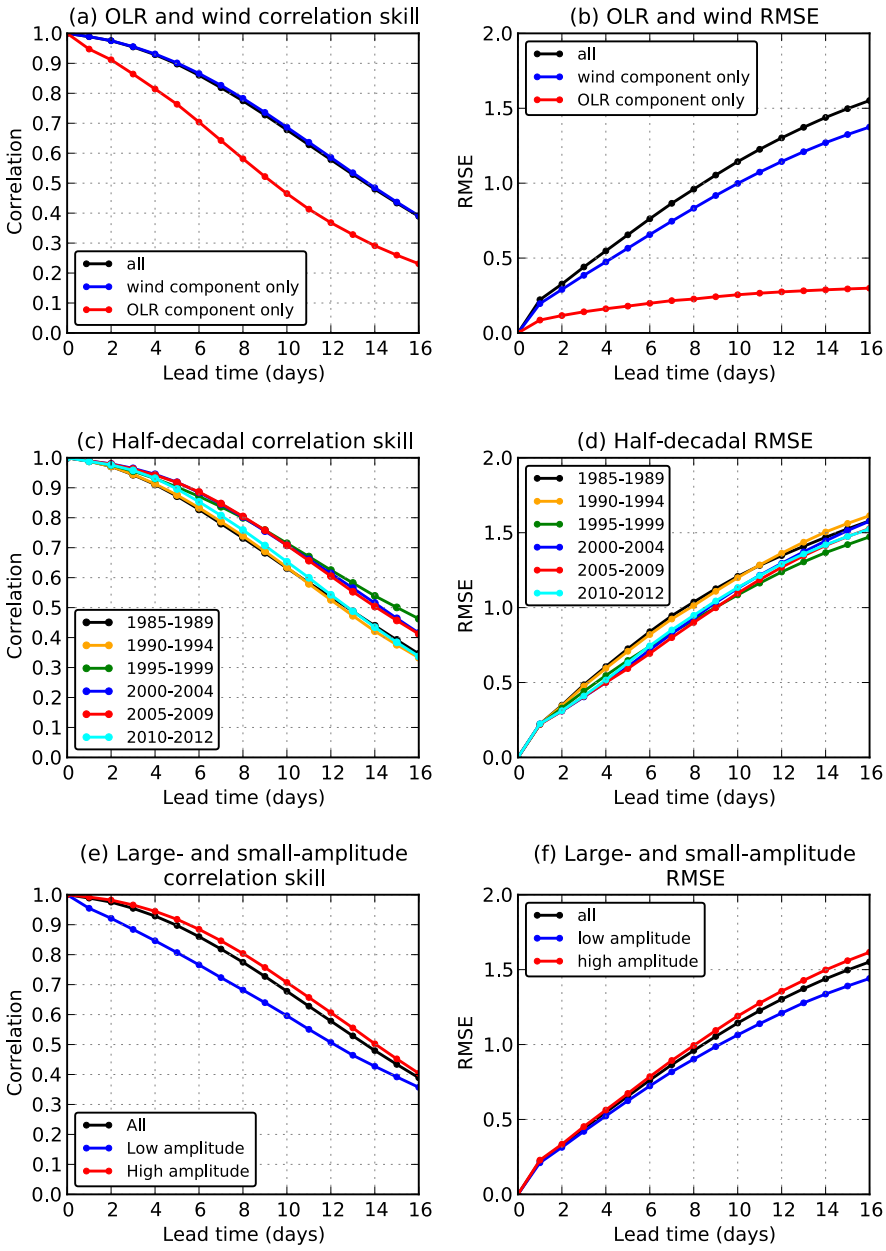


Figure 11: Continuous ranked probability skill scores (CRPSS) of ensemble reforecasts of the MJO relative to an unconditional climatological distribution (red line) and relative to a lagged regression model using current and recent analyzed RMM values as predictors (blue line). The green line shows the "perfect-model" skill, when one forecast member is used as a surrogate for the verification and the remaining 10 members are used to generate the probabilities.



764

765

766 **Figure 12**: Correlation skill (panels (a), (c), and (e)) and RMSE (panels (b), (d), and

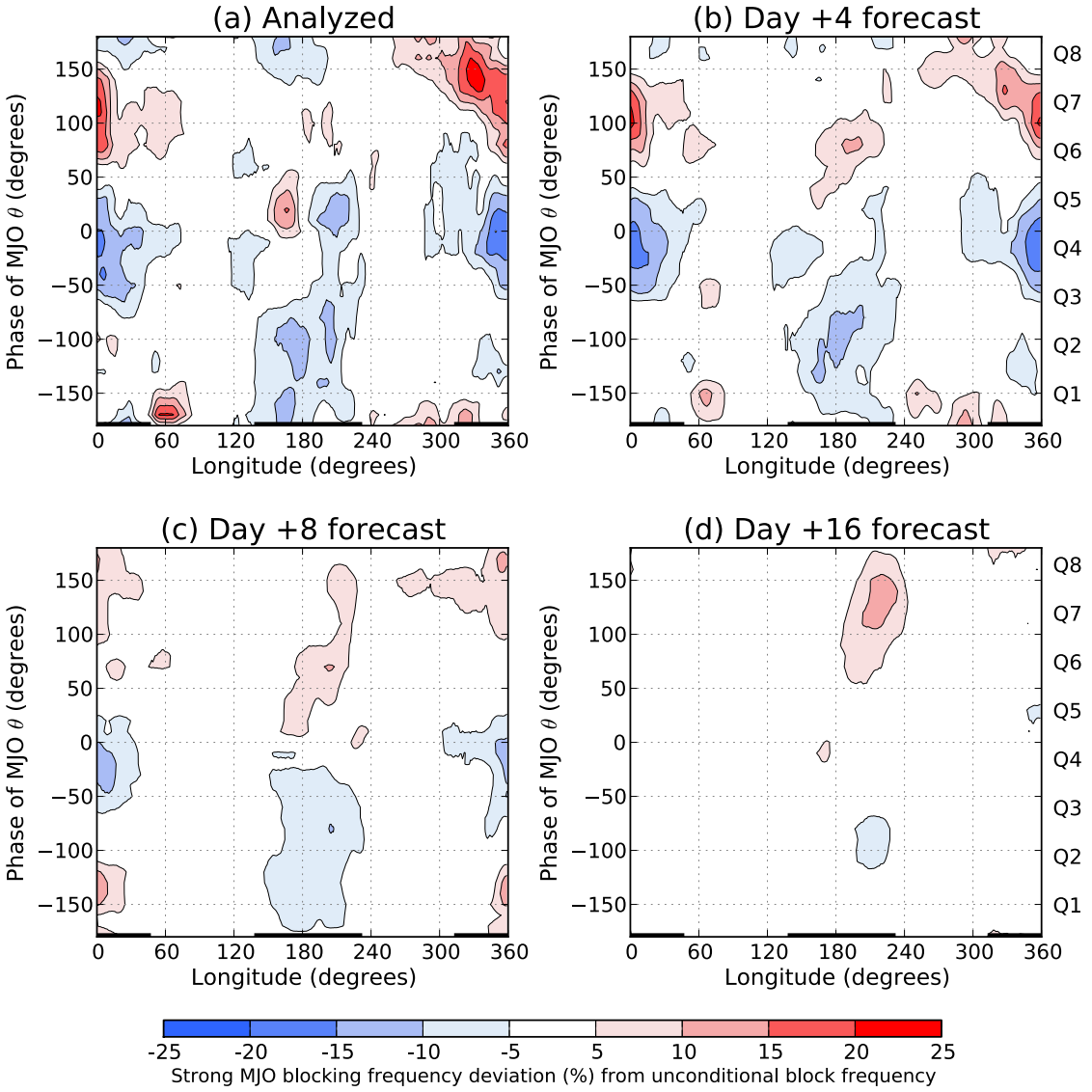
767 (f)) for MJO forecasts. Panels (a) and (b) show skill and RMSE for total and for

768 individual wind and OLR components of RMM. Panels (c) and (d) show overall skill

769 and RMSE for half-decadal periods. Panels (e) and (f) show skill for cases with

770 initial large and small amplitude, as defined in the text.

771



772

773 **Figure 13**: Change in blocking frequency as a function of longitude for when

774 analyzed (panel a) or +4, +8, or +16 day forecast (panels b, c, and d, respectively)

775 has RMM of phase theta + / - 22.5 degrees. Quadrants of MJO diagram as marked on

776 Fig. 6 are noted on the right-hand side of the plots. The ranges the Euro-Atlantic and

777 Pacific Sectors are noted with heavy black lines at the bottom of each panel.